使用程式語言R輕鬆完成試驗設計分析 (四)摺疊設計

農試所作物組 楊滿霞

一、前言

本篇說明使用R函數進行摺疊設計資料之分析，作者自行撰寫摺疊設計的R函數，其函數名稱為NESTED，該函數引數設計及分析結果顯示與前面三篇「使用程式語言R輕鬆完成試驗設計分析 」系列文章類似；其中，變分分析(analysis of variance, ANOVA)以及多重比較最小顯著差異性測驗 (least significant difference, LSD)分析結果，依據使用者回饋意見優化，以利使用者於發表時可以正確且快速完成分析結果整理。

二、摺疊設計簡介

當複因子試驗中，兩因子具有階層或者先後之分，且無法作完全處理組合者，例如動植物雜交後代品系間比較，在農業上可做各雜交組合力的比較或者在某一組合下那個品系較優良；本篇範例為昆蟲品種下不同亞種斑點數之比較，由於在S_1品種下之亞種，不可能同時為S_2品種下之亞種，反之亦然，此種類型之資料稱為摺疊設計。由於因子間不會有完全處理組合，所以在摺疊設計中因子間不會有交感效應存在。特別注意，在變方分析中，上層因子F值以下層因子機差均方為分母，下層因子如顯著，應固定上層因子後，再以LSD檢定下層各變級差異性，延續先前例子，即固定S_1品種後，再比較S_2品種下各亞種之差異性，S_2品種亦然，此部分與完全隨機設計、隨機完全區集設計與裂區設計分析過程較不相同之處 (圖一)。

三、下載NESTED函數

使用NESTED函數前，必須先下載與安裝程式語言R才能順利使用。完成R安裝後，可到農業試驗所官方網站內下載NESTED函數及範例，下載路徑為「農業試驗所>本所簡介>研究單位>作物組>下載專區」，下載「摺疊設計NESTED函數.zip」，內含NESTED.RData及R_InputData資料夾，打開NESTED.RData後，在R console視窗，輸入NESTED後按下enter鍵，即可出現NESTED函數原始程式碼。

四、NESTED函數簡介

R_InputData資料夾內含NESTED _Rprogram.R、f2nested.csv共2個檔案。NESTED_Rprogram.R為NESTED函數的撰寫範例，範例檔案為f2nested.csv。NESTED函數包含5個引數

作　　者：楊滿霞副研究員
連絡電話：04-23317128

(argument)，分別為WD、Filename、X、Up及Dm。以範例NESTED(WD= "d:/R_InputData" ,Filename= "f2nested.csv" ,X=3,Up=1,Dm=2)為說明：

(一)WD為「待分析檔案放置的路徑」，此範例為WD= "d:/R_InputData" (或者WD= 'd:/R_InputData')，等號「=」代表指定的意思，等號後面填寫待分析之csv檔案放置路徑，路徑需用引號 ' ' 或者 " " 標示，此範例代表待分析檔案放在D槽的R_InputData資料夾內，建議使用者可在D槽下建立R_InputData資料夾，並將待分析檔案放入該資料夾內，則不用變動程式，或者請使用者依自己電腦放置待分析檔案位置，調整指定路徑內容，路徑名稱盡量避免出現中文字，以免出現不可預期的錯誤。

(二)Filename為「待分析檔案名稱」，檔案名稱需含副檔名，且檔案必需為csv檔，本範例檔案名稱為f2nested.csv，檔名前後需用引號' '或者" "標示，此範例寫法為Filename=" f2nested.csv" (或者Filename= 'f2nested.csv')。

(三)X為「觀測值」在待分析檔案中的「第幾行」，限制指定的行數只能為1行，此範例寫法為X=3，即f2nested.csv中第3行為觀測值。

(四)Up為「上層因子」在待分析檔案內的「第幾行」，限制指定的行數只能為1行，此範例寫法為Up=1，即f2nested.csv中第1行為上層因子。

(五)Dm為「下層因子」在待分析檔案內的「第幾行」，限制指定的行數只能為1行，此範例寫法為Dm=2，代表f2nested.csv第2行為下層因子。

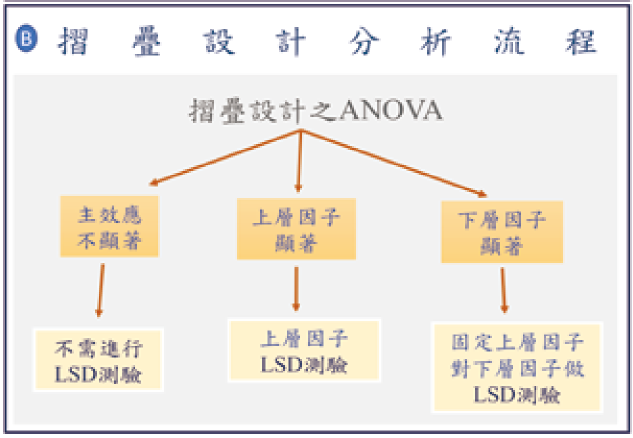
以上為5個引數應填入相對應資訊的說明，填寫在NESTED ()內後並以半形逗號「,」隔開，即完成NESTED函數的撰寫(圖二)。

五、NESTED函數分析結果

有關NESTED函數分析結果以範例NESTED(WD= "d:/R_

A

	摺疊設計		複因子設計	
上層因子	S_1	S_2	S_1	S_2
下層因子	S_{11}	S_{21}	S_{11}	S_{11}
	S_{12}	S_{21}	S_{22}	S_{22}
	S_{13}	S_{21}	S_{33}	S_{33}
處理組合	S_1S_{11}	S_2S_{21}	S_1S_{11}	S_2S_{11}
	S_1S_{12}	S_2S_{22}	S_1S_{22}	S_2S_{22}
	S_1S_{13}	S_2S_{23}	S_1S_{33}	S_2S_{33}



圖一、(A)摺疊設計與複因子設計處理組合比較表；(B)摺疊設計分析流程。

InputData" ,Filename="f2nested.csv" ,X=3,Up=1,Dm=2)做說明。執行該程式後，會出現安裝agricolae套件(該套件功能為分析LSD)的國家連線選擇列表，請選擇「Taiwan」再點選「確定」後，即在R console視窗即出現分析結果，游標點選該視窗後。可使用滑鼠滾珠前後查看分析結果內容，資料分析結果分為3大部分Data Input、ANOVA result和LSD result，說明如下：

(一)Data Input顯示「待分析檔案內的資料簡介」，描述檔案內有22個觀測值共3個變數，也會顯示變數資料的類型，例如「Insect」為chr代表字串，「SubSpecies」為chr代表字串，「Spot」為int代表整數。

(二)ANOVA result顯示「代號設定與變方分析表」，首先將每個變數設定為一個代號，此例觀測值「Spot」代號為x，上層因子「Insect」代號為a，下層因子「SubSpecies」代號為b，

```
NESTED(WD= "d:/R_InputData",Filename= "f2nested.csv",X=3,Up=1,Dm=2)
```

圖二、NESTED函數R程式範例寫法。

A

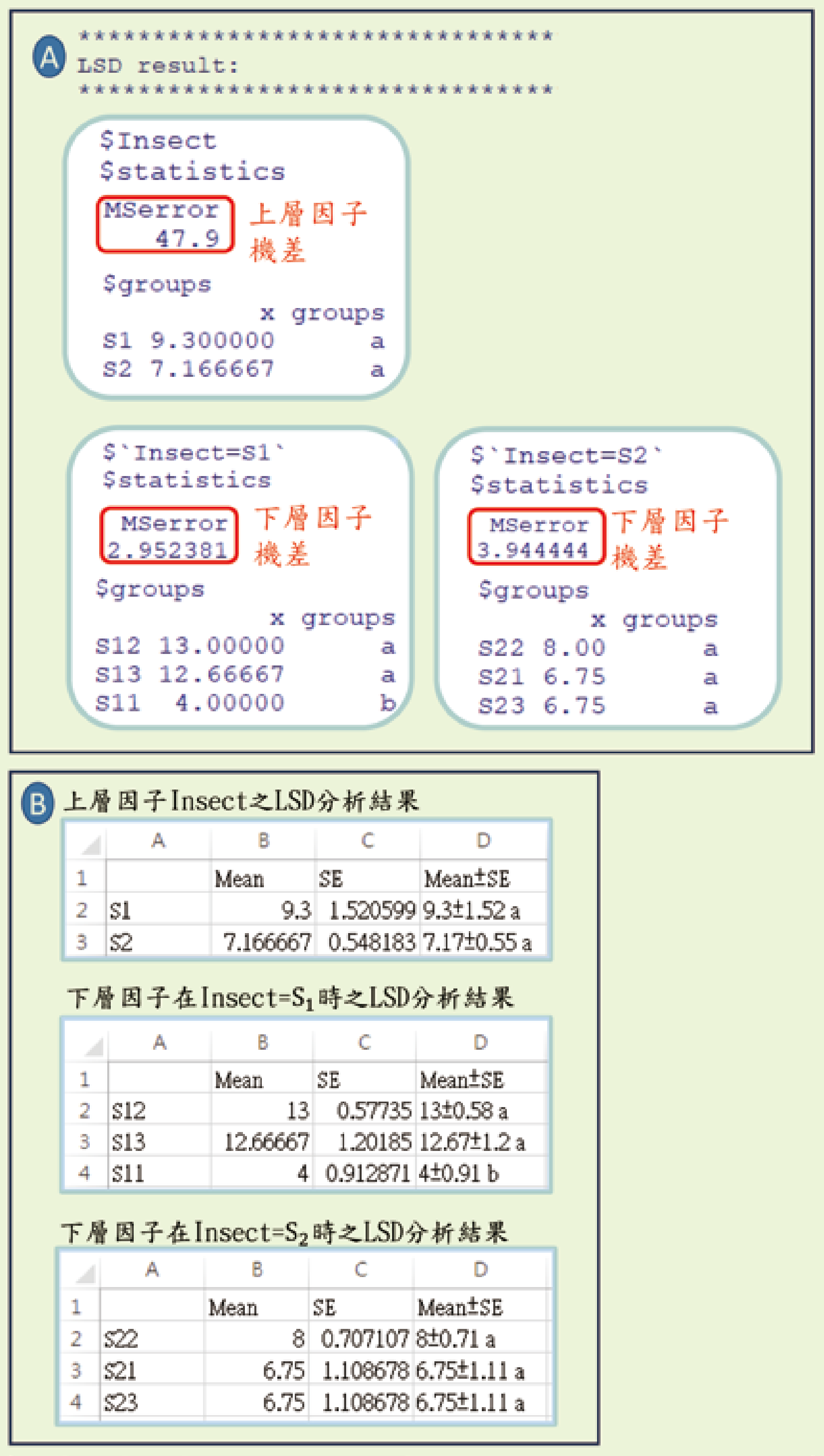
```
*************************************************
ANOVA result:
*************************************************

x= Spot
a= Insect
b= SubSpecies

_________________________________________________

            Df  Sum Sq Mean Sq F value    Pr(>F)
a            1  24.824  24.824  0.5183    0.5114
a:b          4 191.600  47.900 13.6451 5.012e-05 ***
Residuals   16  56.167   3.510
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

上層因子不顯著
下層因子顯著

B

	A	B	C	D	E	F
1		Df	Sum Sq	Mean Sq	F value	Pr(>F)
2	a	1	24.82424	24.82424	0.518251	0.511409
3	a:b	4	191.6	47.9	13.6451	5.01E-05
4	Residuals	16	56.16667	3.510417	NA	NA

圖三、(A) NESTED函數範例在R console視窗顯示ANOVA結果；(B)NESTED函數指定路徑下(以範例來說，為D/:R_InputData/)自動存取一ANOVA分析結果檔案。



A

```
********************************
LSD result:
********************************
$Insect
$statistics
MSerror
   47.9
$groups
           x groups
S1 9.300000      a
S2 7.166667      a

$`Insect=S1`
$statistics
 MSerror
2.952381
$groups
            x groups
S12 13.00000      a
S13 12.66667      a
S11  4.00000      b

$`Insect=S2`
$statistics
 MSerror
3.944444
$groups
        x groups
S22 8.00      a
S21 6.75      a
S23 6.75      a
```


B 上層因子Insect之LSD分析結果

	A	B	C	D
1		Mean	SE	Mean±SE
2	S1	9.3	1.520599	9.3±1.52 a
3	S2	7.166667	0.548183	7.17±0.55 a

下層因子在Insect=S_1時之LSD分析結果

	A	B	C	D
1		Mean	SE	Mean±SE
2	S12	13	0.57735	13±0.58 a
3	S13	12.66667	1.20185	12.67±1.2 a
4	S11	4	0.912871	4±0.91 b

下層因子在Insect=S_2時之LSD分析結果

	A	B	C	D
1		Mean	SE	Mean±SE
2	S22	8	0.707107	8±0.71 a
3	S21	6.75	1.108678	6.75±1.11 a
4	S23	6.75	1.108678	6.75±1.11 a

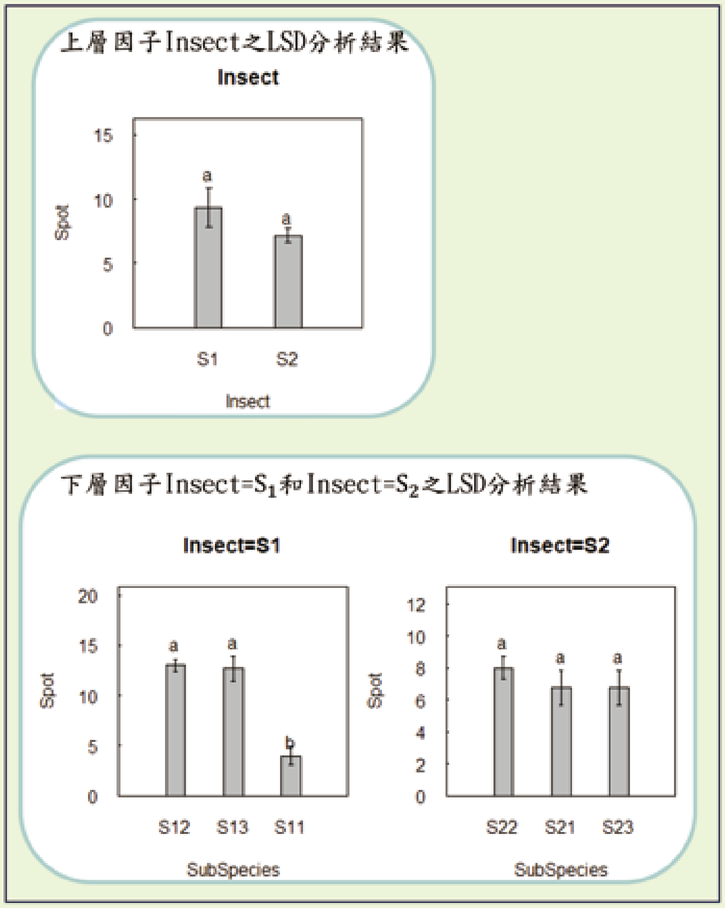
圖四、(A) NESTED函數範例在R console視窗顯示LSD結果；(B) NESTED函數指定路徑下(以範例來說，為D/:R_InputData/)分別自動存取LSD分析結果檔案。

接下來兩部分的ANOVA result和LSD result之內容，變數皆以前述代號顯示。使用者可依變方分析表Pr(>F)欄位檢視處理因子是否顯著，如出現「*」及「**」分別代表達0.05和0.01顯著水準，則須做LSD分析(圖三)。

(三)LSD result依序顯示上層及下層因子分析結果，第一行標示為進行LSD分析之因子名稱，在$group內見到處理LSD分析的結果，其中第一行為處理的代號，第二行為處理平均值，第三行為LSD兩兩比較結果，每個LSD分析結果會以一行底線隔開(圖四)。以NESTED(WD="d:/R_InputData",Filename=

"f2nested.csv" ,X=3,Up=1,Dm=2)分析為例，LSD result可在第一列得知目前這是分析哪一個因子，依序看到上層因子「Insect」、及下層因子SubSpecies在「Insect=S_1」和「Insect=S_2」的分析結果。有關下層因子SubSpecies部分，'Insect=S_1'及在$parameters的name.t欄位為b (SubSpecies)，代表固定Insect因子為S_1時分析b (SubSpecies) 因子的LSD結果。

(四)R Graphics視窗顯示「LSD多重比較標繪圖」，可以按壓page up 或page down檢視不同效應的多重比較結果（圖五），NESTED函數針對每個LSD結果提供使用者誤差線為「標準誤差」之柱狀圖參考，並於柱狀圖上方提供多重比較字母法之結果，使用者可在該視窗點選右鍵複製或儲存圖片。



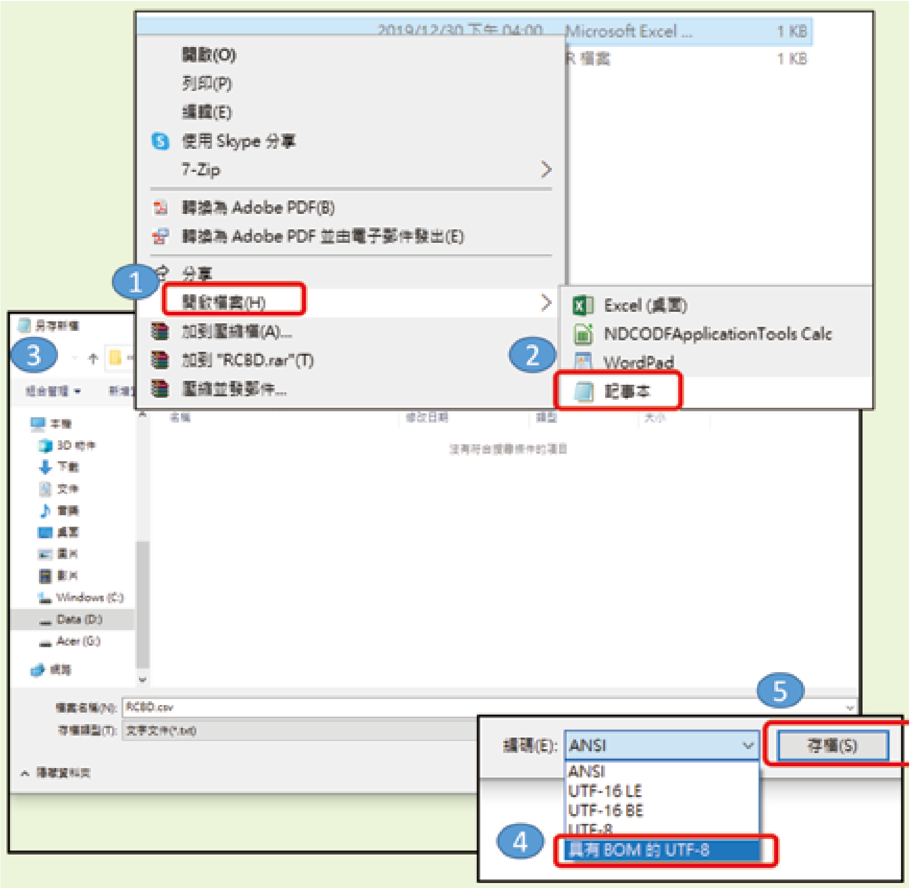
圖五、NESTED函數範例之LSD標繪圖。

六、結語

摺疊設計進行ANOVA分析時上層因子的F值算式需以下層因子機差均方作為分母，但是SASEG(statistical analysis system enterprise guide)軟體產出的ANOVA報表，上層因子F值係以機差做為分母計算而來，尚需使用者自行重新計算正確的主區F值與其對應的P值；此外，上層因子主效應之LSD分析在SASEG軟體中需使用者需特別指定使用機差為「下層因子機差均方」，否則會得到不正確的LSD分析結果。倘若使用NESTED數則不必擔心上述問題，提供正確的ANOVA與LSD分析報表（圖三與圖四），實現撰寫一行NESTED函數，即可正確完成摺疊設計分析。

前三篇系列文章R函數之ANOVA及LSD分析結果皆呈現於R Console視窗，使用者需複製到

EXCEL表格另外進行格式整理，其中LSD的標準差尚需使用者自行估算；作者特別依據使用者回饋意見，在本函數撰寫時優化結果呈現，以本篇範例為例，將ANOVA及LSD分析結果另存於NESTED函數使用者引數指定之路徑下(以範例來說，為D/:R_InputData/)，以檔名為ANOVA.csv儲存ANOVA表格；以檔名LSD_Insect.csv，LSD_Insect=S1.csv及LSD_Insect=S2.csv儲存LSD分析結果，檔案內為各因子變級之平均值、標準差及多重比較字母法標示，可有效提升分析結果整理效率（圖三與圖四）；前述之csv檔案如有亂碼情形，可先以「記事本」方式開啟後，再另存檔案以編碼「具有BOM的UTF-8」存檔，即可正常開啟csv檔案(圖六)。最後特別提醒使用者，在ANOVA分析前，資料須符合常態方布，使用NESTED函數分析前，仍請確認資料是否符合常態分布，以免獲得不正確的分析結果。



圖六、調整csv檔案編碼流程。

七、參考文獻

楊滿霞(2022)。使用程式語言R輕鬆完成試驗設計分析 (一)完全隨機設計。技術服務季刊。131:P32~P37。

楊滿霞(2022)。使用程式語言R輕鬆完成試驗設計分析(二)隨機完全區集設計。技術服務季刊。132:P32~P37。

楊滿霞(2023)。使用程式語言R輕鬆完成試驗設計分析(三)裂區設計。技術服務季刊。134:P25~P29。